Journal of Organizational Computing and Electronic Commerce

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/hoce20

On Latency of E-Commerce Platforms

Marcus Basalla, Johannes Schneider, Martin Luksik, Roope Jaakonmäki & Jan Vom Brocke

JOURNAL OF ORGANIZATIONAL COMPUTING AND ELECTRONIC COMMERCE
2021, VOL. 31, NO. 1, 1–17
https://doi.org/10.1080/10919392.2021.1882240

Taylor & Francis
Taylor & Francis Group

Check for updates

ARTICLE

On Latency of E-Commerce Platforms

Marcus Basalla, Johannes Schneider, Martin Luksik, Roope Jaakonmäki, and Jan Vom Brocke

Institute of Information Systems, University of Liechtenstein, Vaduz, Liechtenstein

ABSTRACT

For e-commerce retailers, even small increases in waiting times have been found to lead to user dissatisfaction and losses in sales. This paper investigates how contextual and behavioral factors influence the impact of latency on the website performance of a real e-commerce platform. The results show an increased impact of latency for users with faster navigation speed and mobile users. They further show a long-term effect of latency on customer returns and confirm previous findings that show a lower impact of latency for users that are familiar with a website. These results provide a case for real-time behavioral analytics by classifying customers based on navigational speed. It further points out two ways to increase revenue: (i) by prioritization of customers and (ii) by optimization of web-pages toward lower latency.

KEYWORDS

Web analytics; latency; e-commerce; navigational speed; big-data; real-time analytics

1. Introductions

Network delay (also called latency in computer network engineering) is often listed as a Key Performance Indicator of a website (Fu et al. 2012). Latency in our context is broader since it encompasses also processing time of the user's browser and transferring data of a webpage. Latency from a user perspective is also referred to as page load time. It begins when the navigation starts, i.e. a website is requested by clicking on a link or entering a URL is completed. It ends when the webpage has finished loading in the user's browser. As such, latency depends on Internet speed, access device and software, and computational needs of the web-page.

Although both Internet and device technologies have improved significantly, latency is still an important factor for web-based companies (Akamai 2017; Google 2017b, 2017a), especially e-commerce platforms (Stadnik and Nowak 2018). First, the effects of faster hardware technology are often offset by more complex and larger websites (Brutlag 2009). And second, even small changes of latency in the order of fractions of a second can have a significant impact on web page usage (Akamai 2017; Google 2017b, 2017a). With a worldwide e-commerce revenue of 1.58 trillion US dollars in 2018 and an estimated revenue growth of 15% per year (Statista 2019), this means that latency has a vast economic impact. Therefore, it is highly important to study the impact of latency on e-commerce in more depth.

Despite the fact, that the validity of old studies (Dellaert and Kahn 1999; Galletta et al. 2004; Marshak and Levy 2003; Nah 2004) might be put into question due to technological changes, there is only limited research on latency and the factors interacting with latency ((Galletta et al. 2006; Hong, Hess, and Hardin 2013) to name two examples). This is especially surprising considering the increase in economic importance of e-commerce since the publication of these studies.

Furthermore, existing studies on latency (Brutlag 2009; Chen, Subramanian, and Toyama 2009; Nah 2004; Stadnik and Nowak 2018; Yu et al. 2020) have an essential shortcoming. It was discovered during discussions with two practitioners responsible for corporate web pages obtaining millions of views yearly. Existing studies are too narrow in their assessment since customers might complete

CONTACT Marcus Basalla marcus.basalla@uni.li University Liechtenstein, Fürst-Franz-Josef-Strasse, 9490 Vaduz, Liechtenstein.

a transaction abandoned due to latency later when latency is either lower or perceived to be less inconvenient.

Moreover, novel opportunities for analytics are not sufficiently leveraged in existing work. Especially existing website usage data has only partially been utilized, despite offering a wealth of information. Among the information routinely recorded by services like Google Analytics is the number of previous website-visits of a user and the device type used for accessing the webpage. The former gives a direct indication of how familiar a user is with a website. This factor could only be indirectly measured in previous studies, for example by providing users with more or less familiar topics or structures of websites (Galletta et al. 2004). In addition, existing data provides currently unused opportunities to derive user-specific features like their navigation speed. Thus, our work aims at addressing the following question:

2. Are familiarity, mobile or non-mobile context and visitors' navigational speed moderating latency?

To answer our question, we analyze a dataset consisting of more than 100.000 sessions of a professional e-commerce store of an international manufacturer of construction equipment. This methodology contrasts with most existing research investigating the influence of latency, which is based on laboratory experiments with relatively few participants over a short time span. We also investigate questions such as the likelihood to return after abandoning a transaction, which is difficult to observe in a laboratory setting. Such challenges that arise in a laboratory setting were also confirmed by (Galletta et al. 2004): "[Intentions] might be difficult to assess when using artificial sites in a laboratory". While we deem laboratory studies highly valuable, we also acknowledge the strengths of studies in a real-world setting.

The increased use of mobile devices can be seen as a major technological development. It is surprising that differences between mobile and non-mobile users with respect to sensitivity to latency have not yet been analyzed scientifically. In particular, since it is known that mobile users behave differently, and websites are commonly designed especially for mobile devices.

In addition to contextual factors like the device used to access a website users' behavioral attributes likely also influence their sensitivity to latency. We therefore investigate the impact of users' navigational speed on latency. For illustration, a person in a hurry seeking to conduct a purchase might only be willing to spend little time on a page, and therefore be more likely to abandon the page if a transaction is prolonged by latency, compared to a person who is more at ease and spends more time on each page to gather information.

Taking the varying impact of latency on users into account allows practitioners to further personalization their e-commerce services. In particular, our results emphasize the importance of website optimization, and the potential to add a further dimension to customer segmentation in web-analytics, based on their sensitivity to latency, using novel attributes such as navigational speed or existing attributes such as page familiarity examined in a real-world e-commerce setting.

In the remainder of this paper, we first give an overview of existing research on latency. We then use existing theories and findings, as well as psychological studies related to moderating factors, to derive a set of hypotheses. These hypotheses are tested on a real-world dataset collected on an online retail platform. In the end, we discuss the implications and limitations of our findings.

3. Related work

This study focuses on time perception and the related factors in an online environment. Previous studies show that time perception is highly subjective and context dependent. However, it also adheres to certain fundamental principles (Egger et al. 2012). Nah (2004) investigated visitors' tolerance to waiting time in an online environment by measuring the time it took users until they abandoned opening a non-working hyperlink. Users experiencing shorter waiting times tended to explore the

website more. Their results point to a tolerable waiting time of two seconds. In addition, they could show that information on the estimated loading time leads to more tolerance for latency. These results are consistent with previous evaluations, showing that delays negatively affect users' evaluation of websites (Dellaert and Kahn 1999). Chen, Subramanian, and Toyama (2009) identified a threshold of ten seconds of waiting time as the crucial moment at which a user gets distracted from their original task. Waiting times above 60 seconds even led users to fully abandon their tasks. Recent research has also discussed how distraction impacts users' behavior during waiting times (Hong, Hess, and Hardin 2013; Lee, Chen, and Hess 2017). For instance, Hong, Hess, and & Hardin (2013) found that visual content can make the wait feel either longer or shorter, depending on the circumstances. More specifically visual content leads to short waiting times (ten seconds) being perceived as longer, and long waiting times (45 seconds) being perceived as shorter.

The increased speed of the Internet and browsing devices since the time of many of these studies have been conducted might have influenced users' sensitivity to latency, as the new generation of Internet users are ever more impatient and expect faster response times (Arapakis, Bai, and Cambazoglu 2014). Arapakis, Bai, and Cambazoglu (2014) investigated the influence of latency on visitors' browsing experience. They found a threshold line of one second, above which visitors noticed a delay with high probability. Website visitors were observed to perform mouse clicks as a result of long waiting times. An experiment by Brutlag (2009) showed, that increased latency in search responses generally led users to conduct fewer searches. An increase of latency from 100 to just 400 milliseconds alone is associated with a decrease of 0.2 to 0.6% in search queries.

While there is little doubt that latency has an impact on users' online behavior, the impact of moderating factors is less well-understood. Ryan and Valverde (2003) conducted a literature review on the effects of latency and factors that influence this effect. Their analysis of 21 papers highlighted that users with more experience in navigating the internet are less tolerant of waiting times. However, in some cases where users already invested time into finishing a certain task, they were more likely to finish that task despite higher latency (Bhatti, Bouch, and Kuchinsky 2000). Other factors that influenced the effects of latency were user's general patience, as well as the information they had about the expected waiting time (Ryan and Valverde 2003).

Galletta et al. (2004) measured the influence of waiting time on users' task performance, attitude, and behavioral intentions in an experiment with 196 participants. Participants had to retrieve information from a mock-up online store. Afterward they completed a survey evaluating their attitude and behavioral intentions. Increased latency led to a clear decrease in all three measured categories. The effect was stronger on a mock-up store that lacked familiar product categories. A follow-up study focused on interaction effects between latency, site breadth, and content familiarity (Galletta et al. 2006). It identified that deeper, more unfamiliar, and slower websites led to a decrease in attitude, performance, and the intention to return to the store. Although users interacted with an online store, these studies did not investigate the influence of latency on e-commerce performance measures like conversion rate. Further these studies only observe users over a short time range. They neglect the possibility that customers might complete a transaction canceled due to latency later when latency is either lower or perceived to be less inconvenient. That is, in Galletta et al. (2006) participants had to endure latency without any option to retry at other times. Also, users were questioned immediately after the experiment about their intention to return, rather than measuring if they returned.

Recently, Yu et al. (2020) investigated the relationship between (web) response times and user experience in mobile applications using 80 participants in a laboratory setting. Their findings are aligned with prior work stating that user experience decreases with growing response times. They also investigated gender and network environment as moderating factors.

Website performance is one key feature in determining user satisfaction in e-commerce, and a factor that affects the decision to purchase a product or not (Poggi et al. 2014). Web visitors leave online sessions in most cases due to a combination of long waiting times, the web's bad performance, and personal impatience. The early drop in e-commerce visitors has an impact on the conversion rate and results in a loss of revenue (Marshak and Levy 2003). A study on an e-commerce website showed,

that users redirected to a high-performance website were 15% more likely to complete a purchase (Nygren, Sitaraman, and Sun 2010). They were also 9% less likely to abandon the site after viewing just one page. According to a recent practitioner report on online retail performance, optimal load times for peak conversions ranged from 1.8 to 2.7 seconds (Akamai 2017). This report was based on data of 10 billion user visits, and it suggests that even 100 milliseconds slower than average page load times decrease the conversion rate by 2.4–7.1%.

Stadnik and Nowak (2018) used Google analytics data to investigate the influence of latency on conversion in an online store. They also provide empirical evidence that conversion rate decreases with latency. They report on average latency of countries and their conversion rates, but do not discuss moderating factors. Similar investigations based on large datasets of e-commerce websites were also performed to investigate the impact of decreased shipping times (Fisher, Gallino, and Xu 2019) and the promised delivery speed Chiu et al., (2020).

Improvements in technology also allow for wide-spread collection of data in the web domain. For example, mouse tracking allows to infer customer preferences in an E-commerce setting (Schneider et al. 2017). Unfortunately, mouse tracking and other forms of data collection also raise privacy concerns, in the light of the General Data Protection Regulation (GDPR) in the European Union.

Though the fundamental primitive of users browsing a web-page during online shopping has not changed during the last twenty years, the online shopping experience has changed considerably. Today's web technology allows for more interactivity, higher quality and more realistic visualizations, improved virtual try-ons (Kim and Forsythe 2008) resulting also in novel quality estimates (Jahromi, Delaney, and Hines 2020). Recent research in artificial intelligence has also contributed notable innovations for E-commerce. For example, deep learning related live assistance based on chatbots (Cui et al. 2020) and recommender systems (Fusco et al. 2019) have improved considerably. Live chat assistance by humans (McLean et al. 2020) or, potentially even by chat bots, can positively impact user trust and purchase intentions. Response times (and therefore latency) is also relevant in the context of such live assistants (Gnewuch et al. 2018).

4. Hypothesis development

The laboratory studies by Galletta et al. (2004) and Nah (2004) both showed that different forms of familiarity impact visitor's sensitivity to latency. Galletta et al. (2004) showed, that visitors were less sensitive to latency on a website that was organized by familiar categories. In the study by Nah (2004), visitors to web pages showed lower tolerance to waiting times before clicking a link, if they experienced large latencies on earlier pages. In our study, familiarity with a website covers both its content as well as loading times. Therefore, in our study, we refer more to familiarity with respect to prior interaction.

Furthermore, we focus on an e-commerce setting, where trust and satisfaction are important factors impacting customers' purchase decisions. In addition, trust and satisfaction also influence the long-term relationship with customers and loyalty toward website vendors Kim, Ferrin, and Rao (2009); Pereira, de Fátima Salgueiro, and Rita (2016). These findings indicate that a loyal regular customer should be more likely to purchase a product or service in the future than a first-time visitor. Regular customers have developed some expectations toward the website performance due to their past perception of page loading times. Customers with only a few previous visits, on the other hand, have no expectations due to prior experiences with the website but base their judgment more on experiences with other sites. This suggests that regular customers should be more willing to accept high latencies, i.e. latency sensitivity decreases with more visits.

Online customers intending to conduct a purchase typically follow a two-stage process (Häubl and Trifts 2000). First, customers investigate a big set of candidate products on more web pages. Afterward, they reduce to a smaller set of alternatives that meet their needs the best. From an efficiency perspective, it is reasonable to abandon a candidate product early in the shopping process due to small inconveniences, while the number of choices is still large, before making a significant effort on collecting information and analyzing the product in more detail. That is to say, customers are more

likely to be sensitive to inconveniences such as latency early in the shopping process. This reasoning is most obvious under the assumption of bounded rationality, which suggests that a user might not investigate all options carefully due to resource constraints (Simon 1990). To make decisions, a user likely employs simple and fast heuristics in decision making to reduce the set of alternatives. This implies that sensitivity to latency decreases across sessions. Some reasons might increase sensitivity to latency, e.g. related to the tolerance level of a user. Users might accept a level of latency above their expected or desired level a few times, but eventually become frustrated and look for alternatives. Still, overall, we posit:

H1a: The number of previous visits reduces the negative impact of latency on conversion.

H1b: The number of previous visits reduces the negative impact of latency on the return likelihood.

Website visitors are commonly segmented into mobile users (e.g. those using a tablet or smartphone) and non-mobile users (e.g. those using a desktop PC or notebook). Over half of all new Internet connections are coming from mobile devices, which has shifted the focus of optimizing website performance only for desktop user's more toward mobile users (Akamai 2017; Gardner 2011). Optimizing latency on mobile devices has its own challenges, as they often have reduced bandwidth and increased latency due to mobile networks, and reduced processing power on mobile devices (Gardner 2011).

Several studies could show that the search behavior of web users differs between mobile and desktop users. This is true for search topics (Church et al. 2008) and search intentions (Song et al. 2013). In addition, it could be shown, that desktop users are willing to invest more time to investigate search results and finding the best possible search outcome. Mobile users, on the other hand, have more concrete and focused queries (Kamvar et al. 2009). Further, mobile users were shown to pay less attention to the display compared to desktop users (Hannak et al. 2014) and service quality for mobile devices is known to fluctuate considerably across time and location (Skocir et al. 2014). These results suggest that desktop users are more patient toward latency compared to mobile users. This is additionally supported by the finding that mobile users are more likely to abandon sessions without clicks (Li and Huffman 2009). An online retail performance report by the cloud service provider Akamai (2017) states that mobile phone and tablet users on average have lower conversation rates than desktop users, and they are more sensitive toward small increases in latency. In addition, the impact of latency on the bounce rate and session duration is stronger for mobile users than desktop users. Based on the previous findings, we formulate the following hypothesis:

H2a: The negative impact of latency on conversion is stronger for mobile users.

H2b: The negative impact of latency on the return likelihood is stronger for mobile users.

Users navigation speed dependent on the experience with a concrete web site, e.g., evidence for learning effects reveals that users spend less time per session the more they visit the site (Johnson, Bellman, and Lohse 2003). More precisely, it seems likely that navigational speed increases with the number of interactions with the page. That is, familiarity with the page might be an important factor. Higher navigational speed might also be a manifestation of being in a hurry, and as such having limited time available to complete a transaction. Indications that the transaction is prolonged, e.g., due to observed latency, seem to make it more likely that the transaction is not completed. In extreme cases, such as in a clinical setting for diagnosing attention deficit hyperactivity disorder (ADHD) (National Institutes of Health 2016) being impatient as well as having difficulty waiting for desired things has been associated with constantly moving and dashing around. Based on these findings, we formulate the following hypothesis:

H3a: The negative impact of latency on conversion is stronger for users with higher navigation speed.

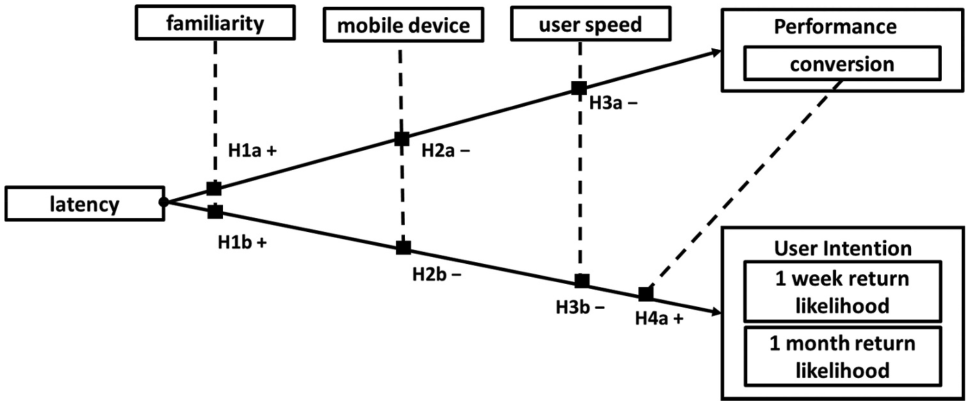

Figure 1. Research model. solid lines indicate direct effects of latency. Dashed lines indicate interaction effects moderating the impact of latency. A "+" sign indicates a positive impact of the independent variable or its interaction with latency on the dependent variables.

H3b: The negative impact of latency on the return likelihood is stronger for users with higher navigation speed.

The idea for the underlying question of the following hypothesis originates from discussions with a corporate website manager. While he believed that latency decreased the chances of completion of a transaction, he said that the company's customers are loyal and therefore they would probably return later to complete the transaction. Thus, latency possibly only delays the completion of a transaction. In the literature, the idea of "retrying" is commonly found in queuing systems (Wang, Na, and Jiang 2010). Prior work has also measured behavioral intentions since they strongly correlate with subsequent behavior (Sheppard, Hartwick, and Warshaw 1988). In particular, in a noncommercial setting Galletta et al. (2004) assessed the intention to return to a website as well as the intention to recommend the webpage dependent on latency. They showed that both decrease with higher latency. Galletta et al. (2004) only measured users' intention to return shortly after the experience with the website using a questionnaire cautioning themselves that their (laboratory) setup has issues. They further did not account for the case of users returning to complete an unfinished task. Thus, prior work does not provide compelling evidence. Given our access to real-world data over a long period, we can measure directly whether a user returned to the website or not, based on actual observed behavior. Therefore, we pose the following hypothesis:

H4: Users that complete a purchase have a higher return likelihood than users who stopped in the middle of a purchase.

All our hypotheses are put in context in the research model in Figure 1.

5. Methodology

We test our hypotheses using a data set from an online retailer recorded by Google Analytics. We use regression analysis to test the influence of the independent variables on the dependent variables.

5.1. Measures/variables

Dependent variables: The measurement of a website's performance often depends on its purpose. As web-based companies follow different business models, they have different Key Performance

Table 1. Distribution of the binary model variables.

	False	True
Conversion	89.35%	10.65%
Returns within 1 week	69.49%	30.51%
Returns within 1 month	62.22%	37.78%
Mobile device	62.29%	37.71%

Indicators (KPIs) (Booth and Jansen 2010). One of the main KPIs for online retail is conversion rate, i.e. the percentage of users who end their visit with a purchase. Since selling products is the main goal of an online retail store and therefore conversation rate is generally the most relevant indicator, we choose this KPI as our dependent variable. For single user sessions, conversion is a binary variable, expressing whether a transaction takes place or not.

Independent variables: Our data contains measures for latency. For this study, we used the time it took from initiating a page view to the page being loaded in the browser as the measure of latency. This value takes both the server-side and the client-side latency into account. In other words, it directly measures how long a user has to wait for a webpage to load. We average the page loading time over all pages visited on a session. To identify fast users (H3), the time until the first action after loading a page for each user was recorded. We consider the first action after the page is loaded to ensure that the user speed variable is not correlated with the latency variable. Users were split into a fast and a slow group at the median, i.e. the faster half of all users, were considered fast. Google Analytics uses the individual user id to count how often a user has visited the website.[1] "According to this number, sessions were assigned to four ordinal groups. Groups were assigned by splitting the data at the quantiles Q25 (2), the Q50 (18), and the Q75 (137) quantile. (H1).

Our data allows differentiating between three device types: desktop, mobile phone, and tablet. To investigate the impact of using a mobile device (H2), the use of a mobile device was interpreted as a binary variable. Both mobile phones and tablets were considered as mobile devices, contrasting them from desktop computers.

We also investigate the mean probability to return within a short time frame (H4). To this end, a binary variable was added. For each session, this variable encodes whether the same customer started a new session within a week of this session's start time. The choice of the interval for return, e.g. one week in our setting, represents a compromise. A (too) long-time period is likely to include many sessions, where a customer might return for a different transaction rather than completing a (suspended) transaction. A short time period is likely to exclude many customers that return after the proclaimed period has elapsed to finish the transaction. The value of one week was chosen loosely based on Ebbinghaus's forgetting curve, stating that retention (of memorized lists) after a week is below 25% (Murre and Dros 2015). To increase the robustness of our results and take longer time windows into account, we added a second binary variable that encodes, whether the customer of a given session started a new session within one month. Table 1 summarizes the distribution of all binary variables except user speed, as this is split at the median.

5.2. Model specification

We test our hypothesis by using multivariate regression models. Thereby we take two factors into account. First, we calculate the p-value for each variable in the model to test whether their influence is significant. Second, we use the R2 value as a measure of how well the models fit the data. In an initial data sighting, we observed, that the effect of latency is highly non-linear, which is aligned with prior findings (Stadnik and Nowak 2018). Small changes in low latency values show a much stronger

[1]Every time a known user visits the website, this count is updated. The count is therefore not influenced by cleaning the data. In other words, if a certain user's previous sessions are not in the current data set, the number of previous visits still takes all their previous visits into account since the analytics system was set up.

influence on the dependent variables as the same changes for higher latency values. We also find that the distribution of the latency variable is strongly skewed to the right. Therefore, a logarithmic transform was applied to the latency values, but no additional parameters were introduced. Figure 2 shows the distribution of latency before and after the logarithmic transformation. The decision can also be justified based on the observed improvements in the R2 values of all models.

Conversion differentiates between two cases, buy or no-buy, which makes it a binary variable. Whether a user returns within a given time is also a binary variable. Therefore, probit regression is adequate to model how the independent variables influence each of the dependent variables. Although our dataset is unbalanced regarding conversions, we do not apply any resampling. This is a valid decision for large datasets as, according to (Crone and Finlay 2012), unbalanced classes do not negatively influence logistic regression, while under-sampling the data to a fraction of its original size might lead to a significant loss of information. As our data set and distribution of positive and negative samples are comparable to the data sets investigated in (Crone and Finlay 2012), we can assume a similar negative impact of under sampling.

5.3. Data

To test the hypotheses, real-world data from an online store of a company producing and selling construction equipment was analyzed. The company sells a large set of products of one brand in the range of a few dollars to more than a thousand dollars. Their operation spans the business-to-business (B2B) as well as the business-to-customer (B2C) domain. The data was collected by the company using Google Analytics. We focused on a time window of several months starting from the 01.01.2018, which gave an original dataset of 371240 individual sessions, as the latency was only recorded for a small percentage of randomly selected sessions[2]. A session here describes one visit to the online shop by a certain user. During one session, the user can navigate the website by visiting different pages.

The data was cleaned by removing all sessions with missing values. We preferred the removal of data to imputation since we found it difficult to find adequate values. The Web Analytics Association standardizes the handling of some missing values, e.g. the visit length is set to zero in case one value such as start or ending time is missing (Burby, Jason, Angie Brown, and WAA Standards Committee 2007). We found that (such) data imputations lead to distortions in our case. While the median latency was approximately 300 milliseconds, the data contained sessions with several seconds loading times. As we were interested in the influence of latency under normal operating conditions, we further removed sessions with extreme latency values. As the distribution of latency values was highly skewed

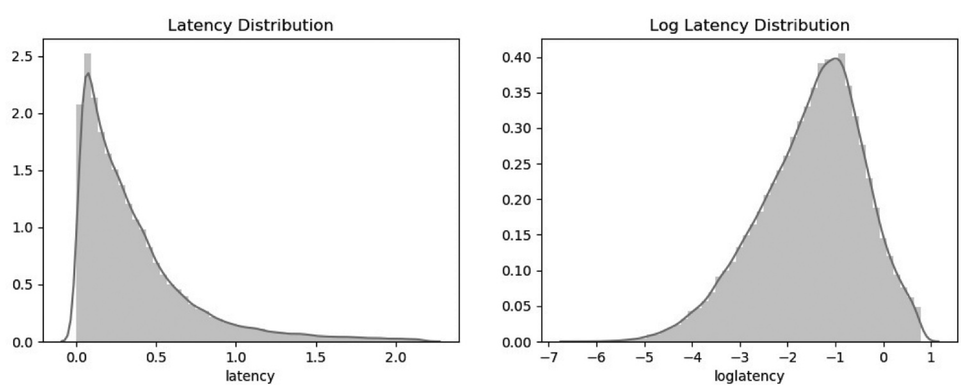

Figure 2. Distribution of latency before and after log transform.

[2]Due to confidentiality we cannot disclose exact numbers.

to the right, we decided to remove the sessions with the highest two percent of latency values. All remaining data had latency values below three seconds.

The dependent variable (number of conversions) was directly recorded in combination with a timestamp, the device type (desktop, mobile, or tablet), and country of origin for each visit, as well as how often the visitor had accessed the online store before (visit number). Individual customers were identified by the unique visitor ID (Burby, Jason, Angie Brown, and WAA Standards Committee (2007). We had no detailed information on customers such as their company, age, or gender.

6. Results

In column (3) of Table 2, we can see a significant impact of the interaction term between loglatency and familiarity on conversions. As a robustness test, we performed a fixed effects regression adding a fixed effect for individual users. The fixed effects allow us to take unobserved individual user effects like the baseline speed of their internet connection, as well as their socioeconomic status into account. As we are investigating a binary dependent variable, we use the estimated fixed effects probit model proposed by Stammann, Heiß, and McFadden (2016).

From column (4) in Table 2, we can see, that even taken the fixed effects into account, the interaction between familiarity and loglatency at $p < .05$. Thus, Hypothesis 1a is supported.

In Table 3 we see a significant impact of the interaction between loglatency and familiarity on users returning within 1 week (column 3) and 1 month (column 6) at $p < .001$, supporting hypothesis H1b. We see that an increase in log latency by 1 decreases the chance of conversion by −0.047%, while a higher familiarity by one category reduces this negative effect on average by 0.0013 percentage points or 1.21% based on the overall conversion rate of 10.72%. The effect is even stronger for the likelihood

Table 2. Regression models for hypothesis H1a with conversion as outcome.

Hypothesis	H1a linear (1)	H1a probit (2)	H1a probit (3)	H1a probit (4)
intercept	0.0257 ***	−1.6999 ***	−1.7697 ***	
	(0.0032)	(0.0146)	(0.0208)	
log latency	−0.0471 ***	−0.2335 ***	−0.2738 ***	−0.5273 **
	(0.0018)	(0.0059)	(0.0103)	(0.0298)
familiarity	0.0089 ***	0.0409 ***	0.0855 ***	0.1033 **
	(0.0017)	(0.0058)	(0.0108)	(0.0349)
log latency × familiarity	0.0013		0.0247 ***	0.0328 *
	(0.0009)		(0.0051)	(0.0141)
Individual fixed effects				Yes
R2	0.0272	0.038	0.0386	0.0812
Observations	67058	67058	67058	19558

Significant *** at $p < .001$; ** at $p < .01$; * at $p < .05$

Table 3. Regression models for hypothesis H1b with return likelihood as outcome.

Hypothesis	H1b linear 1 Week (1)	H1b probit 1 Week (2)	H1b probit 1 Week (3)	H1b linear 1 Month (4)	H1b probit 1 Month (5)	H1b probit 1 Month (6)
intercept	0.2380 ***	−0.6525 ***	−0.7116 ***	0.3132 ***	−0.4059 ***	−0.4878 ***
	(0.0047)	(0.0106)	(0.0142)	(0.0050)	(0.0101)	(0.0135)
log latency	0.0045	0.0582 ***	0.0161 *	−0.0001	0.0581 ***	0.0002
	(0.00271)	(0.0047)	(0.0081)	(0.0029)	(0.0046)	(0.0077)
familiarity	0.0660 ***	0.1497 ***	0.1880 ***	0.0661 ***	0.1203 ***	0.1747 ***
	(0.0026)	(0.0045)	(0.0076)	(0.0028)	(0.0043)	(0.0074)
log latency × familiarity	0.0101 ***		0.0257 ***	0.0141 ***		0.0361 ***
	(0.0012)		(0.0041)	(0.0015)		(0.0039)
R2	0.0178	0.0143	0.0147	0.0139	0.0096	0.0106
Observations	67058	67058	67058	67058	67058	67058

Significant *** at $p < .001$; ** at $p < .01$; * at $p < .05$

Table 4. Regression models for hypothesis H2a with conversion as outcome.

Hypothesis	H2a linear (1)	H2a probit (2)	H2a probit (3)	H2a probit (4)
intercept	0.0411 ***	−1.6016 ***	−1.6377 ***	
	(0.0036)	(0.0155)	(0.0227)	
log latency	−0.0475 ***	−0.2278 ***	−0.2491 ***	−0.5156 ***
	(0.0020)	(0.00609)	(0.0111)	(0.0307)
familiarity	0.0088 **	0.0419 ***	0.0855 ***	0.1044 **
	(0.0017)	(0.0058)	(0.0110)	(0.0350)
log latency × familiarity	0.0011		0.0241 ***	0.0338 *
	(0.0009)		(0.0052)	(0.0141)
mobile	−0.0344 ***	−0.2735 ***	−0.3745 ***	−0.2142 ***
	(0.0040)	(0.0145)	(0.0270)	(0.0550)
mobile × log latency	0.0063 **		−0.0576 ***	−0.0370
	(0.0022)		(0.0128)	(0.0247)
Individual fixed effects				Yes
R2	0.0320	0.0460	0.0462	0.0829
Observations	67058	67058	67058	19558

Significant *** at $p < .001$; ** at $p < .01$; * at $p < .05$

to return, at 3.16% of the overall likelihood to return within one week (31.82%) and a 3.62% of the overall likelihood to return within 1 month (38.95%).

The regression results in column (3) of Table 4 show, that the use of mobile devices increases the negative impact of latency on the dependent variable with $p < .01$. However, this effect is no longer present in the model including fixed effects. Therefore, H2a is only partially supported. Columns (3) and (6) of Table 5 show a significant negative impact of mobile devices on users returning within one week equivalent to −2.30% of the overall return likelihood and one month equivalent to −2,89%of the overall return likelihood respectively, supporting hypothesis H2b.

We cannot find a significant effect of the interaction between latency and user speed on conversion (Table 6 columns (3) and (4)), therefore hypothesis H3a is not supported. However, we can see a significant impact of the interaction on users returning within one week and one month (Table 7). As this effect is not negative Hypothesis H3b is also not supported.

The regression results in Table 8 show a significant positive effect of the interaction between loglatency and conversion on users returning within one week (3.57% of overall return likelihood) and one month (7.56% of overall return likelihood), supporting hypothesis H4.

We also performed a goodness-of-fit test for the models shown in Table 2 to 8. That is, we performed a likelihood ratio test comparing a model with the intercept only to the proposed model

Table 5. Regression models for hypothesis H2b with return likelihood as outcome.

Hypothesis	H2b linear 1 Week (1)	H2b probit 1 Week (2)	H2b probit 1 Week (3)	H2b linear 1 Month (4)	H2b probit 1 Month (5)	H2b probit 1 Month (6)
intercept	0.2298 ***	−0.6867 ***	−0.7329***	0.3128 ***	−0.4242 ***	−0.4882 ***
	(0.0054)	(0.0115)	(0.0162)	(0.0057)	(0.0111)	(0.0155)
log latency	0.0060 *	0.0552 ***	0.0215 *	0.0036	0.0564 ***	0.0103
	(0.0030)	(0.0047)	(0.0091)	(0.0032)	(0.0046)	(0.0087)
familiarity	0.0660 ***	0.1494 ***	0.1880 ***	0.0660 ***	0.1201***	0.1744 ***
	(0.0026)	(0.0045)	(0.00760	(0.0028)	(0.0102)	(0.0074)
log latency × familiarity	0.0102 ***		0.0260 ***	0.0141 ***		0.0362 ***
	(0.0014)		(0.0041)	(0.0015)		(0.0039)
mobile	0.0180 **	0.0793 ***	0.0474 **	0.0007	0.0426***	−0.0001
	(0.0061)	(0.0105)	(0.0175)	(0.0064)	(0.0043)	(0.0170)
mobile × log latency	−0.0074 *		−0.0229 *	−0.0113 **		−0.0306 ***
	(0.0033)		(0.0097)	(0.0035)		(0.0094)
R2	0.0188	0.0149	0.0157	0.0143	0.0098	0.0109
Observations	67058	67058	67058	67058	67058	67058

Significant *** at $p < .001$; ** at $p < .01$; * at $p < .05$

Table 6. Regression models for hypothesis H3a with conversion as outcome.

Hypothesis	H3a linear (1)	H3a probit (2)	H3a probit (3)	H3a probit (4)
intercept	0.0378 ***	−1.6019 ***	−1.6734 ***	
	(0.0038)	(0.0171)	(0.0241)	
log latency	−0.0504 ***	−0.2308 ***	−0.2723 ***	−0.5185 ***
	(0.0023)	(0.0072)	(0.0128)	(0.0336)
familiarity	0.0141 ***	0.0343 ***	0.1073 ***	0.1285 ***
	(0.0021)	(0.0072)	(0.0128)	(0.0378)
log latency × familiarity	0.0056 ***		0.0422 ***	0.0364 *
	(0.0012)		(0.0062)	(0.0154)
fast user	−0.0251 ***	−0.1003 ***	−0.1491 ***	−0.1317 **
	(0.0048)	(0.0162)	(0.0294)	(0.0494)
fast user × log latency	−0.0037		−0.0269	−0.0111
	(0.0027)		(0.0144)	(0.0237)
Individual fixed effects				Yes
Observations	46230	46230	46230	17426
R2	0.0257	0.0094	0.0364	0.1634

Significant *** at $p < .001$; ** at $p < .01$; * at $p < .05$

Table 7. Regression models for hypothesis H3b with return likelihood as outcome.

Hypothesis	H3b linear 1 Week (1)	H3b probit 1 Week (2)	H3b probit 1 Week (3)	H3b linear 1 Month (4)	H3b probit 1 Month (5)	H3b probit 1 Month (6)
intercept	0.2255 ***	−0.7617 ***	−0.7465 ***	0.3008 ***	−0.5082 ***	−0.5291 ***
	(0.006)	(0.0128)	(0.0171)	(0.0058)	(0.0124)	(0.0165)
log latency	0.0108 **	0.0274 ***	0.0373 ***	0.0054	0.0327 ***	0.0155
	(0.0035)	(0.0058)	(0.0105)	(0.0036)	(0.0057)	(0.0101)
familiarity	0.1597 ***	0.4515 ***	0.4279 ***	0.1726 ***	0.4568 ***	0.4571 ***
	(0.0032)	(0.0126)	(0.0094)	(0.0032)	(0.0058)	(0.0095)
log latency × familiarity	−0.0047 **		−0.0163 **	0.0002		0.0020
	(0.0017)		(0.0052)	(0.0018)		(0.0052)
fast user	0.0221 **	0.0305 *	0.0653 **	0.0275 ***	0.0313 *	0.0804 ***
	(0.0072)	(0.0057)	(0.0208)	(0.0073)	(0.0124)	(0.0206)
log latency × fast user	0.0079 *		0.0237 *	0.0117 **		0.0342 **
	(0.0040)		(0.0116)	(0.0040)		(0.0115)
Individual fixed effects						
R2	0.1387	0.1058	0.1060	0.1430	0.1067	0.1068
Observations	46231	46231	46231	46231	46231	46231

Significant *** at $p < .001$; ** at $p < .01$; * at $p < .05$

Table 8. Regression models for hypothesis H4 with return likelihood as outcome.

Hypothesis	H4b linear 1 Week (1)	H4b probit 1 Week (2)	H4b probit 1 Week (3)	H4b linear 1 Month (1)	H4b probit 1 Month (2)	H4b probit 1 Month (3)
intercept	0.2363 ***	−0.6525 ***	−0.7169 ***	0.3082 ***	−0.4076 ***	−0.5014 ***
	(0.0048)	(0.0106)	(0.0144)	(0.0051)	(0.0102)	(0.0137)
log latency	0.0032	0.0582 ***	0.0121	−0.0024	0.0605***	−0.0062
	(0.0028)	(0.0048)	(0.0083)	(0.0029)	(0.0046)	(0.0079)
familiarity	0.0657 ***	0.1497 ***	0.1872 ***	0.0652 ***	0.1200 ***	0.1724 ***
	(0.0026)	(0.0045)	(0.0076)	(0.0028)	(0.0043)	(0.0074)
log latency × familiarity	0.0099 ***		0.0253 ***	0.0137 ***		0.0351 ***
	(0.0014)		(0.0041)	(0.0015)		(0.0039)
Conversion	0.0208	−0.0014	0.0639 ***	0.0775 ***	0.0547 ***	0.2028 ***
	(0.0110)	(0.0168)	(0.0318)	(0.0116)	(0.0162)	(0.0308)
Conversion × familiarity	0.0114 *		0.0345 *	0.0294 ***		0.0777 ***
	(0.0048)		(0.0142)	(0.0051)		(0.0137)
Individual fixed effects						
R2	0.01787	0.0142	0.0148	0.0145	0.0097	0.01104
Observations	67058	67058	67058	67058	67058	67058

Significant *** at $p < .001$; ** at $p < .01$; * at $p < .05$

with dependent variables. We found consistently p-values below 0.001 implying that models with intercept only clearly provide a worse fit to the data than our models.

7. Discussion and implications

7.1. Theoretical and practical implications

Significant changes in web technology and online shopping as witnessed, for instance, by more interactivity, higher quality visualizations, better recommendations and virtual try-ons call existing theory into question. Our findings (H1) strengthens existing theories, e.g. (Galletta et al. 2004), and underline that latency is still an important concern in accordance with other works such as Stadnik and Nowak (2018). Our work suggests further influences on users' perception of and reaction to latency, based on findings from other domains. Research on mobile web user behavior shows that mobile users were less attentive (Hannak et al. 2014), less willing to invest time for investigating search results (Kamvar et al. 2009), and more likely to abandon a session prematurely. This paper could show that a similar effect can be observed regarding latency, namely that mobile users are more strongly affected by latency (H3). By showing that fast users are more strongly affected by latency (H4), our results are linked to theories, that faster user behavior indicates users' impatience, e.g. Ryan and Valverde (2003). Investigating user return behavior (H5) enhances prior theories that exclusively focused on short-lived laboratory experiments by providing a more long-term perspective spanning multiple transactions across multiple days. It shows that latency has a negative long-term impact on visitors' return likelihood. However, this effect is most strongly observed for novel web site users (H2). That is, prior web site users exhibit less willingness to abandon transactions due to latency, which can be explained using interaction costs. Prior users are familiar with the website (probably more than with other sites) therefore experiencing lower interaction costs on the site compared to users not yet familiar with the site. Thus, they have higher incentives to stay on the page despite (increased) latency.

In addition to their theoretical value, our results also lead to practical implications for online retailers. For once, our findings dispelled the belief held by some practitioners that customers abandoning a transaction will (always) return later. Our study re-attested using real-world data and current web-technology that latency impacts conversions permanently. Our discussions with two practitioners responsible for large websites of international companies yielded a strong need for actionable insights with limited investments. We propose two general strategies to take action, namely, mitigate the impact of latency and reduce latency. Latency can be decreased by altering impacting factors, such as web-page design, communication protocols or hardware (Manhas 2013). Measures might be as simple as reordering resource loading (Fainberg et al. 2012), reducing size of multimedia content or, improving server infrastructure, which is often simple in cloud environments. For many webpages, a simple first step might be to maintain two versions of multimedia content that only differ in size, and switch between them based on observed latency. The lower quality version can be automatically generated using down-sizing. In fact, such technology is already commonplace on the web for real-time services such as video streaming or web-conferencing that automatically adjust video resolution and audio quality based on latency. While the web page loading process should be optimized as much as possible in general, our work highlights circumstances and user types that are impacted differently by latency. Measures to improve latency might have adverse effects on other aspects of a webpage such as esthetics, which in turn might also negatively impact conversions. For illustration, for some users it might be better to maintain high quality images and multimedia content at the expense of high latency, while for other users the contrary holds. Thus, to maximize such trade-offs, a personalized treatment and a concise understanding of the impact of latency on a specific user is highly relevant. Our findings suggest that mobile users, i.e. mobile web pages should be addressed first, since mobile users (H3) are more strongly impacted by latency. Furthermore, we observed that visitor's personal attributes like their navigation speed affect their sensitivity to latency (H4). Thus,

high navigational speed might also indicate that latency should be kept lower for a user at the potential expense over worsening other aspects.

Our findings suggest that mobile users, i.e. mobile web pages should be addressed first, since mobile users (H3) are more strongly impacted by latency. Furthermore, we observed that visitor's personal attributes, like their navigation speed affect their sensitivity to latency (H4). Thus, navigational speed might also be a decision criterion whether to prefer latency over other aspects.

Furthermore, in practice, the infrastructure providing webpages might not be able to serve all requests concurrently, as such it might delay some requests more than others. Our findings lend credence to the suggestion of (Poggi et al. 2014, 2009) to dynamically assign server resources to more promising customers. Our findings suggest that mobile users should be addressed first, since mobile users (H3) are more strongly impacted by latency. Furthermore, we observed that visitor's personal attributes, like their navigation speed affect their sensitivity to latency (H4). Furthermore, first time visitor to the website are generally more sensitive to latency (H1). This suggests users with higher navigational speed should be served first to maximize conversions, while people with more prior visits might be kept waiting a bit longer.

Latency might be difficult to change, and it can be that only its effects can be mitigated. That is, for users experiencing latency visual distractions might be used (Hong, Hess, and Hardin 2013). Also, users suffering from high latency might be offered a compensation in the form of discounts. While such measures also induce costs, our work suggests that mobile users, first time visitors and users with high navigational speed experiencing high latency should be addressed first and, probably, with the highest compensation.

7.2. Limitations

There are some limitations to this study with respect to the available data. We only investigate one online shop. Though the shop has a wide range of products from all price ranges, the shop's content is limited to professional equipment by one brand from one type of industry. As most of the sessions originated in Western European or North American countries, the effects might be biased toward these regions and Western cultures though our results do not show strong differences across all regions. As part of the limitation on the data, the conclusions might not be valid to the full extent for both B2B and B2C. While we mentioned that some studies attest little differences between B2B and B2C (Nygren, Sitaraman, and Sun 2010), more investigations would be preferable.

As this is an observational study in an uncontrolled environment, it captures some effects that are likely not prevalent in laboratory studies. This might be a strength with respect to practical relevance but as a limitation with respect to reproducibility and establishing causality. Though given the large number of sessions this might not be the utmost concern, it remains a limitation. For example, we are unaware of a visitor's true intention. This might lead to issues in the selection process of users, when it comes to returning to the website. It could be that users who are particularly interested in a product and not strongly impacted by latency (or users that are generally more tolerant to latency) are those more willing to return. The online shop also provides information on products as such it is not excluded that visitors might not visit the shop with the intention to purchase but rather to investigate product specifications. However, a comparison of the shop investigated in this study and others showed, that the difference in conversion rates was within a reasonable range.

The practical implications of this study are also subject to limitations. First, companies only have limited control over latency, since it is also determined by the customer's device, Internet providers, overall usage of network infrastructure, etc. Thus, measures such as reducing the amount of information transmitted or upgrading the company's computational capabilities, e.g. webservers, might not solve the latency problem completely. Second, latency is only one of several reasons why users abandon a transaction. Therefore, guidelines on how to prioritize customers to maximize website performance should likely also take other factors into account. Further studies might be needed to establish reliable causal links. Additionally, prioritizing customers leads to a violation of the social

norm of "first-come-first-served", which might also upset users and, therefore, this measure on its own might not be effective. It might have to be coupled with compensation schemes for the "injustice" such as discounts for the additional delay.

This study focuses on the conversion rate and customer return likelihood as performance indicators of an e-commerce website. In future work we are planning to investigate the impact of the observed interaction on additional website performance indicators like the overall time spent on the website, the number of individual page views per session, and the overall amount and value of transactions per session.

Furthermore, the data analysis focused on well-interpretable, simple models that are commonly used in scientific work for hypothesis testing. However, in practice, predictive models might be preferable and even needed to apply our findings, although understanding complex models, like deep learning still poses challenges (Meske et al. 2020) and, establishing trust in such models might require more thorough and time-consuming evaluation (Schneider, Handali, and Vom Brocke 2018. However, other fields like Virtual Reality have already provided methodology relating classical scientific analysis, e.g. using regression analysis, of behavioral data and predictive models (Holzwarth et al. 2021).

8. Conclusion

With the prevalence of online retail, the impact of latency on e-commerce is a timely and highly relevant topic for information systems research. Even seemingly low latency has a significant effect on sales as well as on user satisfaction. By reducing latencies user experience could be improved and conversion rates could be increased.

Based on existing literature in the field of human-computer interaction analysis, we identified potential factors, that influence the impact of latency on online retail. We could show that visitors using mobile devices, visitors with high navigational speed, and first-time visitors are more strongly affected by latency. These users are more likely to not finish a transaction in the presence of high latency. In addition, we could show that higher latency has long-term effects that extend beyond a single transaction, i.e. it leads to a decrease in the probability of a visitor returning to the website.

As time perception, in general, is highly subjective and context-dependent but also adheres to fundamental principles (Egger et al. 2012), it is important to understand what these fundamental principles are in the context of e-commerce. Our work sheds some light on moderating factors, but more studies that consider additional factors, such as culture, personal traits, or intentions, are needed.

Notes on contributors

Marcus Basalla is a research assistant and PhD student at the Institute of Information Systems of the University of Liechtenstein. His research focuses on data science and deep learning.

Johannes Schneider is an assistant professor for Data Science at the University of Liechtenstein. His teaching and research focus on data science and machine learning.

Martin Luksik holds a Master's Degree in Information Systems from the University of Liechtenstein. After his Master's he worked as a research assistant at the Institute of Information Systems of the University of Liechtenstein

Roope Jaakonmäki holds a PhD in Enterprise Content Analytics from the University of Liechtenstein. His dissertation focused on scrutinizing the utility of unstructured content by applying a mix of analytics methods, with a focus on unstructured content extracted from social networking sites

Jan vom Brocke is a professor for Information Systems at the University of Liechtenstein. He is the Hilti Endowed Chair of Business Process Management and Director of the Institute of Information Systems.

ORCID

Marcus Basalla http://orcid.org/0000-0003-0609-7336

References

Akamai. 2017. “The state of online retail preformance.” https://www.akamai.com/us/en/multimedia/documents/report/akamai-state-of-online-retail-performance-spring-2017.pdf.

Arapakis, I., X. Bai, and B. B. Cambazoglu. 2014. “Impact of response latency on user behavior in web search.” In Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval - SIGIR ‘14, Gold Coast, Queensland, Australia, eds. S. Geva, A. Trotman, P. Bruza, C. L. A. Clarke, and K. Järvelin, 103–12.

Bhatti, N., A. Bouch, and A. Kuchinsky. 2000. Integrating user-perceived quality into web server design. *Computer Networks* 33 (1–6):1–16. doi:10.1016/S1389-1286(00)00087-6.

Booth, D., and B. J. Jansen. 2010. A review of methodologies for analyzing websites. In *In web technologies: Concepts, methodologies, tools and applications*, ed. A. Tatnall, 145–66. Hershey, PA: Information Science Reference.

Brutlag, J. 2009. “Speed matters for google web search.” http://services.google.com/fh/files/blogs/google_delayexp.pdf.

Burby, Jason, Angie Brown, and WAA Standards Committee. 2007. *Web analytics definitions*. Washington DC: Web Analytics Association.

Chen, J., L. Subramanian, and K. Toyama. 2009. “Web search and browsing behavior under poor connectivity.” Proceedings of the 18th international conference on World wide web, Madrid, Spain.

Church, K., B. Smyth, K. Bradley, and P. Cotter. 2008. “A large scale study of European mobile search behaviour.” Proceedings of the 10th international conference on Human computer interaction with mobile devices and services, Amsterdam, The Netherlands.

Crone, S. F., and S. Finlay. 2012. Instance sampling in credit scoring: An empirical study of sample size and balancing. *International Journal of Forecasting* 28 (1):224–38. doi:10.1016/j.ijforecast.2011.07.006.

Cui, R., T. Sun, L. Zhikun, and J. Golden. 2020. “Sooner or later? Promising delivery speed in online retail.”, Available at SSRN 3563404

Dellaert, B. G. C., and B. E. Kahn. 1999. How tolerable is delay? Consumers’ evaluations of internet web sites after waiting. *Journal of Interactive Marketing* 13 (1):41–54. doi:10.1002/(SICI)1520-6653(199924)13:1<41::AID-DIR4>3.0.CO;2-S.

Egger, S., T. Hossfeld, R. Schatz, and M. Fiedler. 2012. “Waiting times in quality of experience for web based services.” Fourth International Workshop on Quality of Multimedia Experience, Melbourne, Australia.

Fainberg, L., O. Ehrlich, G. Shai, O. Gadish, D. O. B. O. Amitay, and O. Berger (2012). *U.S. Patent No. 8,250,457*. Washington, DC:U.S. Patent and Trademark Office.

Fisher, M. L., S. Gallino, and J. J. Xu. 2019. The value of rapid delivery in omnichannel retailing. *Journal of Marketing Research* 56 (5):732–48. doi:10.1177/0022243719849940.

Fu, Q., J.-G. Lou, Q.-W. Lin, R. Ding, D. Zhang, Y. Zihao, and T. Xie. 2012. “Performance issue diagnosis for online service systems.” IEEE 31st Symposium on Reliable Distributed Systems, NW Washington, DC, United States.

Fusco, F., M. Vlachos, V. Vasileiadis, K. Wardatzky, and J. Schneider. 2019. “RecoNet: An interpretable neural architecture for recommender systems.“ In Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, Macao, China, 2019, Eds. Sarit Kraus, 2343–49.

Galletta, D. F., R. Henry, S. McCoy, and P. Polak. 2004. Web site delays: How tolerant are users. *Journal of the Association for Informatio Systems* 5 (1):1–28. doi:10.17705/1jais.00044.

Galletta, D. F., R. M. Henry, S. McCoy, and P. Polak. 2006. When the wait isn’t so bad: The interacting effects of website delay, familiarity, and breadth. *Information Systems Research* 17 (1):20–37. doi:10.1287/isre.1050.0073.

Gardner, B. S. 2011. Responsive web design: Enriching the user experience. *Sigma Journal: Inside the Digital Ecosystem* 11 (1):13–19.

Gnewuch, U., S. Morana, M. Adam, and A. Maedche 2018. Faster is not always better: Understanding the effect of dynamic response delays in human-chatbot interaction.

Google. 2017a. “Find out how you stack up to new industry benchmarks for mobile page speed.” Accessed April 30, 2019. https://www.thinkwithgoogle.com/_qs/documents/57/mobile-page-speed-new-industry-benchmarks.pdf.

Google. 2017b. “Mobile site speed: Playbook.” Accessed April 30, 2019. https://www.thinkwithgoogle.com/_qs/documents/4290/c676a_Google_MobileSiteSpeed_Playbook_v2.1_digital_4JWkGQT.pdf.

Hannak, A., G. Soeller, D. Lazer, A. Mislove, and C. Wilson. 2014. “Measuring price discrimination and steering on e-commerce web sites.” Proceedings of the 2014 conference on internet measurement conference, New York, NY, United States.

Häubl, G., and V. Trifts. 2000. Consumer decision making in online shopping environments: The effects of interactive decision aids. *Marketing Science* 19 (1):4–21. doi:10.1287/mksc.19.1.4.15178.

Holzwarth, V., J. Schneider, J. Handali, J. Gisler, C. Hirt, A. Kunz, and J. Vom Brocke. 2021. Towards estimating affective states in virtual reality based on behavioral data. In *Virtual reality*.

Hong, W., T. J. Hess, and A. Hardin; A. & Hardin. 2013. When filling the wait makes it feel longer: A paradigm shift perspective for managing online delay. *Mis Quarterly* 37 (2):383–406. doi:10.25300/MISQ/2013/37.2.04.

Jahromi, H. Z., D. T. Delaney, and A. Hines. 2020. Beyond first impressions: Estimating quality of experience for interactive web applications. *IEEE Access* 8:47741–55. doi:10.1109/ACCESS.2020.2979385.

Johnson, E. J., S. Bellman, and G. L. Lohse. 2003. Cognitive lock-in and the power law of practice. *Journal of Marketing* 67 (2):62–75. doi:10.1509/jmkg.67.2.62.18615.

Kamvar, M., M. Kellar, R. Patel, and Y. Xu. 2009. “Computers and iphones and mobile phones, oh my! a logs-based comparison of search users on different devices.” In Proceedings of the 18th international conference on World wide web, New York, NY, United States, 801–10.

Kim, D. J., D. L. Ferrin, and H. R. Rao. 2009. Trust and satisfaction, two stepping stones for successful e-commerce relationships: A longitudinal exploration. *Information Systems Research* 20 (2):237–57. doi:10.1287/isre.1080.0188.

Kim, J., and S. Forsythe. 2008. Adoption of virtual try-on technology for online apparel shopping. *Journal of Interactive Marketing* 22 (2):45–59. doi:10.1002/dir.20113.

Lee, Y. (G.), A. N. K. Chen, and T. Hess. 2017. The online waiting experience: Using temporal information and distractors to make online waits feel shorter. *Journal of the Association for Information Systems* 18 (3):231–63. doi:10.17705/1jais.00452.

Li, J., and S. Huffman, and A. & Tokuda. 2009. “Good abandonment in mobile and PC internet search.” Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, New York, NY, United States.

Manhas, J. 2013. A study of factors affecting websites page loading speed for efficient web performance. *International Journal of Computer Sciences and Engineering* 1 (3):32–35.

Marshak, M., and H. Levy. 2003. Evaluating web user perceived latency using server side measurements. *Computer Communications* 26 (8):872–87. doi:10.1016/S0140-3664(02)00220-7.

McLean, G., K. Osei-Frimpong, A. Wilson, and V. Pitardi. 2020. How live chat assistants drive travel consumers' attitudes, trust and purchase intentions. *International Journal of Contemporary Hospitality Management* 32 (5):1795–812. doi:10.1108/IJCHM-07-2019-0605.

Meske, C., E. Bunde, J. Schneider, and M. Gersch. 2020. Explainable artificial intelligence: objectives, stakeholders, and future research opportunities. *Information Systems Management* 1–11. doi:10.1080/10580530.2020.1849465.

Murre, J. M. J., and J. Dros. 2015. Replication and analysis of Ebbinghaus' forgetting curve. *PloS One* 10 (7):e0120644. doi:10.1371/journal.pone.0120644.

Nah, F. 2004. A study on tolerable waiting time: How long are web users willing to wait? *Behaviour & Information Technology* 23 (4):2212–22. doi:10.1080/01449290410001669914.

National Institutes of Health. 2016. “Attention-Deficit/Hyperactivity Disorder (ADHD): The basics.” Accessed April 26, 2019. https://www.nimh.nih.gov/health/publications/attention-deficit-hyperactivity-disorder-adhd-the-basics/index.shtml.

Nygren, E., R. K. Sitaraman, and J. Sun. 2010. The akamai network: A platform for high-performance internet applications. *ACM SIGOPS Operating Systems Review* 44 (3):2–19. doi:10.1145/1842733.1842736.

Pereira, H. G., M. de Fátima Salgueiro, and P. Rita. 2016. Online purchase determinants of loyalty: The mediating effect of satisfaction in tourism. *Journal of Retailing and Consumer Services* 30:279–91. doi:10.1016/j.jretconser.2016.01.003.

Poggi, N., D. Carrera, R. Gavalda, E. Ayguadé, and J. Torres. 2014. A methodology for the evaluation of high response time on E-commerce users and sales. *Information Systems Frontiers* 16 (5):867–85. doi:10.1007/s10796-012-9387-4.

Poggi, N., J. L. Toni Moreno, R. G. Berral, J. Torres, and J. Torres. 2009. Self-adaptive utility-based web session management. *Computer Networks* 53 (10):1712–21. doi:10.1016/j.comnet.2008.08.022.

Ryan, G., and M. Valverde. 2003. Waiting online: A review and research agenda. *Internet Research* 13 (3):195–205. doi:10.1108/10662240310478213.

Schneider, J., J. P. Handali, and J. Vom Brocke 2018. Increasing trust in (big) data analytics. In *International Conference on Advanced Information Systems Engineering,Tallinn, Estonia,* 70–84.

Schneider, J., M. Weinmann, J. Vom Brocke, and C. Schneider 2017. *Identifying preferences through mouse cursor movements: Preliminary evidence*. Proccedings of the European Conference of Information Systems (ECIS), Guimarães, Portugal.

Sheppard, B. H., J. Hartwick, and P. R. Warshaw. 1988. The theory of reasoned action: A meta-analysis of past research with recommendations for modifications and future research. *Journal of Consumer Research* 15 (3):325–43. doi:10.1086/209170.

Simon, H. A. 1990. Bounded rationality. In *Utility and probability*, 15–18. The Macmillan Press Limited, London.

Skocir, P., D. Katusic, I. Novotni, I. Bojic, and G. Jezic. 2014. “Data rate fluctuations from user perspective in 4G mobile networks.” In 22nd International Conference on Software, Telecommunications and Computer Networks (SoftCOM), Split, Croatia, 180–85.

Song, Y., H. Ma, H. Wang, and K. Wang. 2013. “Exploring and exploiting user search behavior on mobile and tablet devices to improve search relevance.” Proceedings of the 22nd international conference on World Wide Web, Rio de Janeiro, Brazil, 1201–12.

Stadnik, W., and Z. Nowak. 2018. “The impact of web pages' load time on the conversion rate of an E-commerce platform.” In Information Systems Architecture and Technology: Proceedings of 38th International Conference on Information Systems Architecture and Technology – ISAT 2017, Szklarska Poręba, Poland.

Stammann, A., F. Heiß, and D. McFadden (2016): Estimating fixed effects logit models with large panel data. In: Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Online at: http://hdl.handle.net/10419/145837.

Statista. 2019. "eCommerce: Worldwide." Accessed August 18, 2019. https://www.statista.com/outlook/243/100/ecommerce/worldwide.

Wang, K., L. Na, and Z. Jiang. 2010. "Queueing system with impatient customers: A review." Proceedings of 2010 IEEE International Conference on Service Operations and Logistics, and Informatics, QingDao, China, 82–87.

Yu, M., R. Zhou, Z. Cai, C. W. Tan, and H. Wang. 2020. Unravelling the relationship between response time and user experience in mobile applications. *Internet Research* (5):1353–82. doi:10.1108/INTR-05-2019-0223.